Project Document: Project Ideas
Team 1

Cam Carlson
Bo Paxton
Pranav Nikam
Nick Montemarano

Contents

Abstract

Deep learning is a powerful tool which can assist people in their day to day lives. We are exploring various ways to implement machine learning to make the lives of people easier by offering solutions and improvements to common issues in society. Specifically, the two ideas we propose are, a machine learning model which can detect the presence of a wallet in some image. This implementation could be used by CCTV or security systems to assist people in recovering their lost belongings. And second, using a more text based approach, we can classify movie reviews to determine if they are positive or negative. While not as crucial, this idea would allow logistics in the film industry to be more accurate, and enable producers to better understand their audiences' tastes.

Chapter 1

Milestone 1: Project Ideas

1.1 Introduction

While brainstorming potential areas of life to implement deep learning, the level of importance that our project would have was a key factor. While the ideas presented here are not curing any diseases or writing symphonies, the quality of life improvements that can be achieved are quite large.

1.2 Project Idea 1: Lost Wallet Identifier

Given recent advances in object identification in crowded scenarios using deep learning computer vision models, we propose a system that can identify potential lost wallets, using security cameras. Such a system could be highly useful in public places like airports, malls, or transportation areas where lost personal belongings are common. This project aims to build a deep learning model capable of detecting whether or not a wallet is present in an image, potentially contributing to real-time security and object recovery systems.

1.2.1 Applications

Potential applications for this project may include (but are not limited to):

- Automated lost-and-found detection in public spaces

- Surveillance systems for tracking lost objects

- Airport and station security enhancements

- Assistance for individuals with visual impairments

1.2.2   Approaches

Different approaches to this problem may include:

- **Deep Learning**: Using learning algorithms such as the Adam optimizer [2], a model can be trained using some dataset, and tested on some unique validation set. Supervised learning is a potential approach.

- **Data Augmentation**: To increase the robustness of the model, data augmentation techniques such as random cropping, rotation, and flipping can be applied to increase the variety of training samples and prevent overfitting.

1.2.3   Learning Problem

The learning problem is a **binary classification** problem, where the task is to classify an image as either containing a wallet or not. This project may also include an element of object detection if the model needs to locate the wallet within the image.

1.2.4   Inputs and Outputs

- **Inputs**: Image data captured from cameras, potentially including cluttered backgrounds.

- **Outputs**: A binary label indicating whether or not a wallet is present in the image.

1.2.5   Data Sources

Potential data sources may include:

- Public datasets for object detection, such as COCO (Common Objects in Context), which could be modified or supplemented with wallet-specific data.

- Custom image datasets, potentially gathered from publicly available sources or manually curated by collecting images of wallets in various real-world settings.

1.3   Project Idea 2: Film Review Classification

As for the second proposed idea, according to the study *The Influence of Movie Reviews on Consumers* [1], positive reviews have the greatest impact on influencing whether people choose to watch a movie. For filmmakers and producers, this distinction between positive and negative feedback can be a valuable tool in shaping audience perception. Understanding and gauging audience reactions is crucial for success in the film industry.

The problem is to develop a deep learning model that can binary classify a movie review as positive or negative. This classification can help filmmakers and producers understand audience reactions and make informed decisions.

1.3.1 Applications

Potential applications for this project may include (but are not limited to):

- Sentiment analysis for movie reviews

- Audience perception analysis

- Film industry decision-making

- Recommendation systems for movies

1.3.2 Approaches

Different approaches to this problem may include:

- **Machine Learning**: Train a binary classifier using machine learning algorithms such as Gradient Descent using the Adam optimizer [2].

- **Deep Learning**: Utilize deep learning architectures such as Supervised Learning or Unsupervised Learning.

- **Natural Language Processing (NLP)**: Employ NLP techniques such as text preprocessing, tokenization, and feature extraction to analyze the sentiment of movie reviews.

1.3.3 Learning Problem

The learning problem is a **binary classification** problem, where the goal is to classify a movie review as either positive or negative.

1.3.4 Inputs and Outputs

- **Inputs**: Movie reviews in the form of text data

- **Outputs**: Binary classification labels (positive or negative)

Table 1.1: Contributions by team member for Milestone 1.

Team Member	Contribution
Cam Carlson	Project writeup
Pranav Nkam	Project ideas and research
Nick Montemarano	Scheduling and time management
Bo Paxton	Troubleshooting and backup help

1.3.5 Data Sources

Potential data sources could include non-commercial datasets from:

- IMDB

- Rotten Tomatoes

- Kaggle

1.4 Conclusions

We have more interest in working on project 1 due to its practicality however we feel like we have a more confidence and better understanding of working with project 2.

If given a choice, we would do project 1, however we would need extra assistance on what types of datasets we would need, what computer vision model would be the best to work with, and how we should test our model

Remember to cite sources using BiBTeX and add those references to the end of this document! Sample citation [**Samuel59**]

Chapter 2

Milestone 2: Project Selection

In this chapter, you should select your project, formally define it, and propose two approaches, each with a work plan. You should follow the The Heilmeier Catechism [**heilmeier**], which is a set of questions used to evaluate proposed research programs:

HC1 What are you trying to do? Articulate your objectives using absolutely no jargon.

HC2 How is it done today, and what are the limits of current practice?

HC3 What is new in your approach and why do you think it will be successful?

HC4 Who cares? If you are successful, what difference will it make?

HC5 What are the risks?

HC6 How much will it cost?

HC7 How long will it take?

HC8 What are the mid-term and final "exams" to check for success?

Also, remember to update your abstract when you've completed this chapter.

2.1 Introduction

Explain which project you chose, and why. (This may differ from the two you proposed in the previous chapter.) **Include citations.**

2.2 Problem Specification

Formally define your chosen problem, including the following, as subsections[1]:

1. A statement of your project topic (HC1) that includes a formal discussion of the learning problem, including what the inputs and outputs of your trained model will be.

2. Motivation for your topic: Why it is important and interesting (HC3, HC4). If your work is in an application area, be careful to avoid technical jargon from that area that is outside of computer science. If you must use a term, define it as carefully and simply as possible.

3. What resources will you need, including data sets and libraries that need to be installed on `swan` (HC6).

4. Cite at least three references (at least two published journal or conference papers) (HC2).

2.3 Related Work

Here, summarize other approaches to this problem. You should cite each paper and briefly discuss it. **Reviewing related work is critical early on in your project work, because the literature gives ideas on what approaches work and don't work, potential pitfalls, baseline performance to set your expectations, etc.**

2.4 Proposed Method 1: [add name of method and delete brackets]

Propose a step-by-step approach to solve your chosen problem. You should include a precise work plan: What you plan to do, what data sets you will test on (including the size of each data set), how you will preprocess the data, what the steps (pipeline) of this method are (e.g., what NN architecture(s) will be used and how they'll be linked), how you will evaluate performance, what baseline(s) from the literature you will compare against, what is your timeline, etc. (HC5, HC7, HC8). One or more of your references should relate to aspects of this.

---

[1]The four items below are in an enumerated list to make the requirements clear. Your write-up should not do this; simply have paragraphs of text addressing all the required items.

Table 2.1: Contributions by team member for Milestone 2.

Team Member	Contribution
Member1	Contribution1
Member2	Contribution2
Member3	Contribution3

2.5 Proposed Method 2: [add name of method and delete brackets]

Propose a step-by-step approach to solve your chosen problem. You should include a precise work plan: What you plan to do, what data sets you will test on (including the size of each data set), how you will preprocess the data, what the steps (pipeline) of this method are (e.g., what NN architecture(s) will be used and how they'll be linked), how you will evaluate performance, what baseline(s) from the literature you will compare against, what is your timeline, etc. (HC5, HC7, HC8). One or more of your references should relate to aspects of this.

2.6 Conclusions

Sum up, including your opinion on each approach. Also, list any questions that you have for the instructor and TA regarding your project work.

Remember to cite sources using BibTeX and add those references to the end of this document!

It is okay to cite some websites and tutorials (if you first look up how to properly cite them!), but you must also cite some refereed publications from conferences and/or journals.

Finally, in Table 2.1, list each member of your team with a brief summary of that member's contribution to this milestone.

Chapter 3

Milestone 3: Progress Report 1

3.1 Introduction

You will choose one method from Milestone 2 and conduct experiments on it. (Note that you might end up changing your method one or more times before Milestones 4 and 5, depending on your results with it. But for now, you may just focus on a single approach.)

Remind the reader of the problem you're working on, and what your approach is. Briefly summarize what your results are so far, including quantification and **comparison to baseline approach(es)**.

Also, remember to update your abstract when you've completed this chapter.

3.2 Related Work

Here, summarize other work related to yours. You should cite each paper and briefly discuss it, comparing and contrasting it with your approaches and results. **Reviewing related work is critical early on in your project work, because the literature gives ideas on what approaches work and don't work, potential pitfalls, baseline performance to set your expectations, etc.**

3.3 Experimental Setup

Describe the setup of your experiments so far in sufficient detail for the reader to reproduce them. Include data sources, preprocessing used, architectures/other approaches used, hyperparameter values used, performance measures used, baseline(s) from the literature that you compare against, and other relevant items (e.g., cross-validation).

Table 3.1: Contributions by team member for Milestone 3.

Team Member	Contribution
Member1	Contribution1
Member2	Contribution2
Member3	Contribution3

3.4   Experimental Results

Present your results so far using tables, figures, confusion matrices, etc. (whatever is appropriate) for your approach(es) and the baseline(s).

3.5   Discussion

Discuss your experimental results, drawing conclusions that are supported by your experimental results of Section 3.4. Be careful to not draw conclusions that are not supported by the evidence you present!

3.6   Work Plan

Discuss how the results of Section 3.5 will influence your future experiments. Include descriptions of what new things you will try, why you'll try them, and what your timeline is for the rest of the semester.

3.7   Conclusion

Sum up, including your a summary of your results so far (recapitulating that from Section 3.1). Also, describe your plans for future work and list any questions that you have for the instructor and TA regarding your project work.

Remember to cite sources using BiBTeX and add those references to the end of this document!

It is okay to cite some websites and tutorials (if you first look up how to properly cite them!), but you **must** also cite some refereed publications from conferences and/or journals.

Finally, in Table 3.1, list each member of your team with a brief summary of that member's contribution to this milestone.

Chapter 4

Milestone 4: Progress Report 2

4.1 Introduction

Remind the reader of the problem you're working on, and what your approach is. Briefly summarize what your results are so far, including quantification and **comparison to baseline approach(es)**.

Also, remember to update your abstract when you've completed this chapter.

4.2 Related Work

Here, summarize other work related to yours. You should cite each paper and briefly discuss it, comparing and contrasting it with your approaches and results. **Reviewing related work is critical early on in your project work, because the literature gives ideas on what approaches work and don't work, potential pitfalls, baseline performance to set your expectations, etc.**

4.3 Experimental Setup

Describe the setup of your experiments so far in sufficient detail for the reader to reproduce them. Include data sources, preprocessing used, architectures/other approaches used, hyperparameter values used, performance measures used, baseline(s) from the literature that you compare against, and other relevant items (e.g., cross-validation).

Table 4.1: Contributions by team member for Milestone 4.

Team Member	Contribution
Member1	Contribution1
Member2	Contribution2
Member3	Contribution3

4.4 Experimental Results

Present your results so far using tables, figures, confusion matrices, etc. (whatever is appropriate) for your approach(es) and the baseline(s).

4.5 Discussion

Discuss your experimental results, drawing conclusions that are supported by your experimental results of Section 4.4. Be careful to not draw conclusions that are not supported by the evidence you present!

4.6 Work Plan

Discuss how the results of Section 4.5 will influence your future experiments. Include descriptions of what new things you will try, why you'll try them, and what your timeline is for the rest of the semester.

4.7 Conclusion

Sum up, including your a summary of your results so far (recapitulating that from Section 4.1). Also, describe your plans for future work and list any questions that you have for the instructor and TA regarding your project work.

Remember to cite sources using BiBTeX and add those references to the end of this document!

It is okay to cite some websites and tutorials (if you first look up how to properly cite them!), but you must also cite some refereed publications from conferences and/or journals.

Finally, in Table 4.1, list each member of your team with a brief summary of that member's contribution to this milestone.

Chapter 5

Milestone 5: Final Report

5.1   Introduction

Remind the reader of the problem you're working on, and what your approach is. Briefly summarize what your final results are, including quantification and **comparison to baseline approach(es)**.

Also, remember to update your abstract when you've completed this chapter.

5.2   Related Work

Here, summarize other work related to yours. You should cite each paper and briefly discuss it, comparing and contrasting it with your approaches and results. **Reviewing related work is critical early on in your project work, because the literature gives ideas on what approaches work and don't work, potential pitfalls, baseline performance to set your expectations, etc.**

5.3   Experimental Setup

Describe the setup of your experiments in sufficient detail for the reader to reproduce them. Include data sources, preprocessing used, architectures/other approaches used, hyperparameter values used, performance measures used, baseline(s) from the literature that you compare against, and other relevant items (e.g., cross-validation).

5.4   Experimental Results

Present your results using tables, figures, confusion matrices, etc. (whatever is appropriate) for your approach(es) and the baseline(s).

Table 5.1: Contributions by team member for Milestone 5.

Team Member	Contribution
Member1	Contribution1
Member2	Contribution2
Member3	Contribution3

5.5   Discussion

Discuss your experimental results, drawing conclusions that are supported by your experimental results of Section 5.4. Be careful to not draw conclusions that are not supported by the evidence you present!

5.6   Conclusion

Sum up, including your a summary of your results (recapitulating that from Section 5.1). Also, describe possible avenues for future work (should one continue the project) and list any questions that you have for the instructor and TA regarding your project work.

Remember to cite sources using BiBTeX and add those references to the end of this document!

It is okay to cite some websites and tutorials (if you first look up how to properly cite them!), but you must also cite some refereed publications from conferences and/or journals.

Finally, in Table 5.1, list each member of your team with a brief summary of that member's contribution to this milestone.

Appendix A

First Appendix

An appendix is used only if necessary (remove this chapter if you don't use one). It contains supplementary materials/extra details such as extensive experimental results (e.g., hyperparameter search results, where the most interesting ones are in the main text and the rest are dumped here), detailed proofs, etc.

Create as many appendices as needed by adding chapters. All such chapters must be before the bibliography.

main